

**UNITED STATES DISTRICT COURT
DISTRICT OF NEW JERSEY**

IN RE: INSULIN PRICING LITIGATION

**Case 2:23-md-03080
MDL No. 3080**

**JUDGE BRIAN R. MARTINOTTI
JUDGE RUKHSANAH L. SINGH**

DECLARATION OF DOUGLAS FORREST

I, Douglas Forrest, state and declare as follows:

1. I am the Senior Vice President, eDiscovery Analytics & Strategy, at International Litigation Services (“ILS”), based in Irvine, California (www.ilsteam.com). I have been retained as a consultant for the Plaintiffs in this action. The facts stated in this declaration, except as otherwise explicitly noted, are within my own personal knowledge and, if called as a witness to testify, I could and would competently testify to the facts contained in this declaration.

2. I make this declaration in connection with Plaintiffs’ responses to the submissions of Defendants Express Scripts, CVS Caremark, and Optum. In preparing this declaration I reviewed:

- a. CMO 11 (ECF No. 208) (“ESI Order”);
- b. Letter from Thomas P. Scrivo to the Honorable Rukhsanah L. Singh, U.S.M.J. (ECF No. 412) (“Optum February 5 Letter”);
- c. Transcript of Feb. 10, 2025 Hearing, *In Re: Insulin Pricing Litigation*, MDL No. 3080 (D.N.J.) (ECF No. 426) (“Hearing Transcript”);
- d. Email from Steven J. Daroci to Nicholas Suellentrop (February 28, 2025) (“Plaintiffs’ February 28 Email”).
- e. Plaintiffs’ Letter Re Custodian and Search Term Disputes Involving PBM Defendants (February 28, 2025) (ECF No. 439) (“Plaintiffs’ Letter Brief”);

- f. Letter from Jason R. Scherr to the Honorable Rukhsanah L. Singh, U.S.M.J. (February 28, 2025) (ECF No. 436) (“Express Scripts February 28 Letter”);
- g. Letter from John D. Tortorella to the Honorable Rukhsanah L. Singh, U.S.M.J. (February 28, 2025) (ECF No. 437) (“CVS Caremark February 28 Letter”);
- h. Letter from Thomas P. Scrivo to the Honorable Rukhsanah L. Singh, U.S.M.J. (February 28, 2025) (“Optum February 28 Letter”);
- i. Email from Nicholas Suellentrop to Steve Daroci (March 3, 2025) (“CVS Caremark March 3 Email”); and
- j. CVS Caremark’s Hit Counts (Plaintiffs 2.28.25 Terms).pdf (March 3, 2025) (“CVS Hit Report”).

I. QUALIFICATIONS

3. My qualifications are set forth in the Declaration of Douglas Forrest previously submitted in support of Plaintiffs’ Proposed ESI Order (ECF No. 123-2) and are incorporated herein by reference.

II. OPINIONS

Background

4. At the hearing held before Magistrate Judge Singh on February 10, 2025, the Court directed the parties “to continue that meet-and-confer [on search terms and custodians]. If you cannot reach an agreement by February 24, the parties are to propose their respective search terms and custodians by way of letter addressed to [the Court] to be filed by February 28.” Hearing Transcript, p. 10.

5. Two of the PBM Defendants, CVS Caremark and Optum, submitted letters pursuant to the Court’s direction objecting to search terms. (CVS Caremark February 28 Letter and Optum February 28 Letter). I understand that Express Scripts, the third PBM Defendant, has

accepted the Plaintiffs' search terms in connection with its proposed implementation of TAR.
(Express Scripts February 28 Letter).

***Plaintiffs Terms Are Neither Overbroad nor Excessive for
this Complex, Far-reaching, and High-Stakes Litigation***

6. Two of the PBM Defendants — CVS Caremark and Optum — advance burden
claims with respect to Plaintiffs' proposed search terms based upon the number of document hits:

a. **CVS Caremark:**

"Indeed, Plaintiffs February 19 proposal hit on nearly four million documents.... the addition of more custodians will raise these hit counts (perhaps substantially)." CVS February 28 Letter, p. 3.

"Plaintiffs' proposal still hits on over 3.3 million documents with families for only the first 15 custodians listed in CVS's February 28, 2025 letter to the Court (ECF No. 437), not the broader set of 21 custodians that we proposed to the Court, which is obviously not acceptable." CVS March 3 Email.

b. **Optum:**

Plaintiffs' terms "returned millions of hits" within 12 custodians' files and concluded that Plaintiffs' proposal would "pose an undue burden on OptumRx." Optum February 5 Letter, p. 2. *See also* Optum February 28 Letter, pp. 2-3.

7. These claims must be considered in the context of a case of this size and
complexity, as described on page 9 of Plaintiffs' Letter Brief.

8. It is commonplace that millions of documents will be collected and produced in a
case of this size and complexity. For example, ILS¹ has provided hosting in at least four MDLs
and other complex cases where the productions of individual defendants were over 4 million
documents, including one case where an individual defendant produced over 7 million
documents, and another where an individual defendant produced over 9 million documents.

¹ This is based on an analysis that I requested from my ILS colleagues.

9. As these are counts of *produced* documents, the number of documents *collected* by each individual defendant necessarily would have been significantly larger than the counts of produced documents.

10. Plainly put, there is nothing untoward or unusual in the number of documents questioned here by CVS Caremark and Optum. In fact, compared to cases of comparable scope and complexity, Plaintiffs' search terms may even be less broad than they could have been.

11. Moreover, it is important in evaluating CVS Caremark's and Optum's burden claims to note what they haven't claimed in their respective February 28 letters.²

12. While producing responsive documents in the sine qua non of discovery, these PBM Defendants do not dispute that Plaintiffs' search terms will return responsive documents.

13. The PBM Defendants also do not present any specific examples of wide swaths or types of non-responsive documents, which would provide data to enable Plaintiffs to refine their search terms to exclude such document types.

14. CVS Caremark does complain that the set of Plaintiffs' search terms "is not proportional to the needs of the case, is overly burdensome, and is unnecessarily cumulative" (CVS Caremark February 28 Letter, p. 3), but:

- a. CVS Caremark doesn't offer any explanation for why Plaintiffs' search terms would not be proportional, which those search terms plainly would be given the complexity, scope, and stakes of the litigation;

² CVS Caremark does claim that the number of documents "would bring this action largely to a halt for more than one year, if not longer, while documents are reviewed." With all due respect to Sedona 6, and a producing party's prerogatives, this is akin to a California-based witness protesting that it would take them months to travel to a trial in NY as they intended to travel by horseback.

- b. CVS Caremark's claim that Plaintiffs' search terms would be overly burdensome basically boils down to the number of documents with search term hits *collected*, even though, as discussed *supra*, Defendants in other complex cases, including MDLs, have *produced* many more documents as would be the norm in comparable cases of this complexity and scope, with comparable stakes.
- c. CVS Caremark doesn't offer any explanation why Plaintiffs' search terms would be unnecessarily cumulative or offer any specific examples of any documents that would be cumulative.

15. Consideration must also be given to the simple fact that Plaintiffs have had virtually no access to Optum's collection, and access to no more than a small fraction, a little over 5%, of CVS Caremark's collected documents.³ In trying to formulate search terms without such access, Plaintiffs are basically in the same position as the blind men trying to describe an elephant in the famous parable.⁴

16. This consideration is magnified by the basic fact that search terms are an abysmal way to identify a reasonable proportion of responsive documents.

17. The deficiencies of search terms in litigation were exposed in the 1985 publication of the landmark paper by Blair and Maron⁵ that found that only 20% of relevant

³ My calculation: 176,775 (CVS Caremark documents produced) divided by 3,334,995 (Total documents with hits, including Group, from CVS Hit Report).

⁴ See, e.g., *Blind men and an elephant*, Wikipedia, https://en.wikipedia.org/wiki/Blind_men_and_an_elephant (accessed on March 5, 2025).

⁵ David C. Blair & M. E. Maron, An Evaluation of Retrieval Effectiveness for a Full-Text Document-Retrieval System, 28 COMM'CNS ACM 289, 290 (1985).

documents included search terms selected by experienced counsel (who thought that the search terms would have retrieved 75% of relevant documents).

18. Later studies at the 2007 and 2008 TREC Legal Track⁶ confirmed this dismal state of affairs, finding estimated recall rates for search terms of 22% (i.e., 78% of relevant documents missed), and 24%⁷ (i.e., 76% of relevant documents missed). Perhaps worse, with respect to highly relevant documents, the result of the application of a “consensus set” of Boolean search terms in the 2008 Legal Track implied “that, on average per topic, *58% of the ‘highly relevant’ documents were not found* ... indicating that it is not just tangentially relevant documents that are being missed by the negotiated Boolean approach.”⁸

19. Given this horrendous track record, and the consequent high likelihood that a significant proportion of responsive documents will be excluded from production by using search terms, Plaintiffs would be prejudiced were their search terms not accepted.

⁶ The Text Retrieval Conference (TREC), co-sponsored by the NIST Information Technology Laboratory's (ITL)

Retrieval Group was established in 1992. *Celebrating 25 Years of TREC*, available at <https://trec.nist.gov/celebration/25thcelebration.html>. The Legal Track at the Text REtrieval Conference (TREC) was established “to assess the ability of information retrieval techniques to meet the needs of the legal profession for tools and methods capable of helping with the retrieval of electronic business records, principally for use as evidence in civil litigation.” TREC Legal Track, available at <https://trec-legal.umiacs.umd.edu/>.

⁷ Stephen Tomlinson, et al., Overview of the 2007 TREC Legal Track (April 30, 2008), (22% recall), available at

<https://trec.nist.gov/pubs/trec16/papers/LEGAL.OVERVIEW16.pdf>; Douglas W. Oard, et al., Overview of the TREC 2008 Legal Track (November 1, 2008), (finding 24% recall), available at <https://apps.dtic.mil/sti/pdfs/ADA512711.pdf>.

⁸ Oard, *supra*, at p. 39 (emphasis added).

Hit Reports

20. Hit reports, while imperfect, can provide some information useful in search term negotiations, as recognized by the ESI Order which states:

“If a Producing Party elects to use search terms to identify potentially responsive documents and ESI, it shall identify and propose to the Requesting Party an initial list of search terms, custodians, custodial data sources, and non-custodial data sources that are likely to contain responsive documents and ESI, and the Parties will meet and confer regarding those terms and any additional terms, or other requested changes, proposed by the Requesting Party. That meet and confer shall include, when requested, the exchange of information that includes, where relevant, semantic synonyms, code words, acronyms, abbreviations, nonlanguage alphanumeric associational references to relevant ESI, any *hit count reports*, and any other information agreed to by the Parties. To the extent the Parties cannot reach agreement on the application of, or procedures for, any search or filtering processes, the Parties may raise such issues for resolution by the Court or its designee.”

ESI Order, Section V.A. (emphasis added).

21. The imperfections of search term hit reports start with the fact that they offer no information at all with respect to the responsive documents that *don't* include search terms, which are likely to constitute well over half of the responsive documents.

22. Moreover, hit reports do not provide any information about the importance of any document. A smoking gun, i.e., a document that conclusively establishes a critical element of proof, could contain only a single search term which is also contained by many other documents of little or no probative value making that search term a tempting candidate for deletion in search term negotiations.

23. That said, properly comprehensive hit reports can provide some useful metrics.

24. For example, these metrics give some idea of the number of documents that would require review:

- (1) total documents containing hits;
- (2) total documents containing hits, including family members;

(3) total unique documents with hits;

(4) total unique hits with families.

25. But even these metrics are imperfect. For example, the metric unique hits for each search query gives an imperfect measure of how many documents would be removed from the collection if that search query was removed. It's imperfect because if more than one search query is removed, it would remove not only all the documents unique to each (which would be known numbers) but also all the documents hit only by one or more of the search queries being removed. For example, if search query A uniquely hit 100 documents and search query B uniquely hit 100 documents, removing either A or B would remove only 100 documents but removing A and B would remove not only the 200 (100 + 100) documents unique to each term, but also any documents which hit only on A *and* B, without any indication of how many documents – 100, 1,000, 10,000 – that would be.

26. I understand Plaintiffs have requested hits reports from at least certain Defendants but have gotten varying levels of response. For example, on February 28, 2025, Plaintiffs requested that CVS Caremark provide hit reports with the following metrics:

1. Hits + Family: the number of documents that hit on the search term, including their family members
2. Unique Hits: the number of documents that hit on the search term but do not have hits for any other terms
3. Totals for each category above
 - Total documents containing hits
 - Total documents containing hits, including family members
 - Total unique documents with hits

27. On March 3, 2025, CVS Caremark provided a hit report (CVS Hit Report) which contained data for requests 1, 2, and 3.3 as well as data for Total documents with hits, including Group, and Total documents without hits.

28. Plaintiffs included other important requests in their request for hit counts:

1. Please run the final list by your vendor to edit asterisks, parentheses, capitalization, etc. and let us know if any revisions were made.
2. Please confirm how misspellings are accounted for, including whether fuzzy searching is being used to find variants of the search terms and what fuzziness level is being used.
3. Please confirm whether the search strings were run as “aggregated” sets or as individual terms broken out.
4. Please provide a description of the corpus of documents against which the hit reports were run.
 - E.g., Total volume, custodians, sources, date range, whether deduplicated prior to being run, and whether email thread suppressed prior to being run.

Plaintiffs’ February 28, 2025 Email.

29. CVS Caremark ignored requests 2 and 4, and responded only to the first and third requests, stating:

“Certain modifications were made to the terms in order to run them properly through the database. We have attached those modifications here. Further, certain of Plaintiffs’ search string are too long to run together. They were thus broken up as illustrated in the attached.” CVS Caremark March 3 Email.

30. I also understand that Plaintiffs do not know with certainty what search platforms are being used by any Defendant — another important data point to inform any meaningful conferral over search terms — and therefore Plaintiffs cannot evaluate such assertions or ascertain what platform capabilities would be available to address misspellings.

31. Search queries must be platform specific as the available functionality and query operators and the required query syntax will vary from platform to platform. For example, what proximity operators does the platform support – e.g., just w/#n (within #n words) or also such useful operators as /s (within the same sentence) and /p (within the same paragraph).

32. Are regular expressions, a/k/a regexes, supported? For example, using regular expressions “/p” searches could be implemented on a platform that doesn’t have a “/p” operator.

33. Other platform and implementation considerations and the interplay between them may also come into play. For example, if email header information such as Subject is not

included in extracted full text, how will proximity searches succeed if the only mention of a search term used with a proximity operator is in an email's Subject metadata field which isn't indexed for or included in full text searches, and the search term on the other side of the proximity operator is indexed and included in full text searches?

DATED: March 5, 2025

s/ Douglas Forrest

Douglas Forrest

Senior Vice President, eDiscovery
Analytics & Strategy
International Litigation Services, Inc.
dforrest@ilsteam.com